RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/719,024
Source: IFW0
Date Processed by STIC: 11-15-04

ENTERED

IFWO

RAW SEQUENCE LISTING DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:37

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

<110> APPLICANT: University of Kentucky Research Foundation
JONES, Grace
JONES, Davy
<120> TITLE OF INVENTION: MUTANTS AND ASSAY SYSTEM TO IDENTIFY USP/RXR LIGANDS
<130> FILE REFERENCE: 050229-0420
<140> CURRENT APPLICATION NUMBER: 10/719,024
<141> CURRENT FILING DATE: 2003-11-24
<150> PRIOR APPLICATION NUMBER: 60/428,282
<151> PRIOR FILING DATE: 2002-11-22
<160> NUMBER OF SEQ ID NOS: 23
<170> SOFTWARE: PatentIn version 3.3
<210> SEQ ID NO: 1
<211> LENGTH: 2488
<212> TYPE: DNA
<213> ORGANISM: Drosophila melanogaster
<400> SEQUENCE: 1

CP5.6)

aaaaatgtcg acgcgaaaaa aggtatttat tcattagtca gaaagtctgg cattctttgt 60
ttgttggtaa aaagcgcaat tgtttggagg cgagcgaata aagtgcgctg ctccatcggc 120
tcaagattat gtaaatgcag caacgacccc accaacaacg aaactgcaac ctgctccact 180
tggcccaacg gaccaatagc ggacggacgg acacggtggc gttggcaaag tgaaacccca 240
acagagaggc gaaagcgagc caagacacac cacatacaca cgaagagaac gagcaagaag 300
aaaccggtag gcggaggagg cgctgccccc agttcctcca atatacccag caccacatca 360
caagcccagg atggacaact gcgaccagga cgccagcttt cggctgagcc acatcaagga 420
ggaggtcaag ccggacatct cgcagctgaa cgacagcaac aacagcagct tttcgcccaa 480
ggccgagagt cccgtgccct tcatgcaggc catgtccatg gtccacgtgc tgcccggctc 540
caactccgcc agctccaaca acaacagcgc tggagatgcc caaatggcgc aggcgcccaa 600
ttcggctgga ggctctgccg ccgctgcagt ccagcagcag tatccgccta accatccgct 660
gagcggcagc aagcacctct gctctatttg cggggatcgg gccagtggca agcactacgg 720
cgtgtacagc tgtgagggct gcaagggctt ctttaaacgc acagtgcgca aggatctcac 780
atacgcttgc agggagaacc gcaactgcat catagacaag cggcagagga accgctgcca 840
gtactgccgc taccagaagt gcctaacctg cggcatgaag cgcgaagcgg tccaggagga 900
gcgtcaacgc ggcgcccgca atgcggcggg taggctcagc gccagcggag gcggcagtag 960
cggtccaggt tcggtaggcg gatccagctc tcaaggcgga ggaggaggag gcggcgtttc 1020
tggcggaatg ggcagcggca acggttctga tgacttcatg accaatagcg tgtccaggga 1080
tttctcgatc gagcgcatca tagaggccga gcagcgagcg gagacccaat gcggcgatcg 1140
tgcactgacg ttcctgcgcg ttggtcccta ttccacagtc cagccggact acaagggtgc 1200
cgtgtcggcc ctgtgccaag tggtcaacaa acagctcttc cagatggtcg aatacgcgcg 1260
catgatgccg cactttgccc aggtgccgct ggacgaccag gtgattctgc tgaaagccgc 1320
ttggatcgag ctgctcattg cgaacgtggc ctggtgcagc atcgtttcgc tggatgacgg 1380
cggtgccggc ggcgggggcg gtggactagg ccacgatggc tcctttgagc gacgatcacc 1440
gggccttcag ccccagcagc tgttcctcaa ccagagcttc tcgtaccatc gcaacagtgc 1500
gatcaaagcc ggtgtgtcag ccatcttcga ccgcatattg tcggagctga gtgtaaagat 1560
gaagcggctg aatctcgacc gacgcgagct gtcctgcttg aaggccatca tactgtacaa 1620

RAW SEQUENCE LISTING DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:37

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

81 cccggacata cgcgggatca agagccgggc ggagatcgag atgtgccgcg agaaggtgta 1680
83 cgcttgcctg gacgagcact gccgcctgga acatccgggc gacgatggac gctttgcgca 1740
85 actgctgctg cgtctgcccg ctttgcgatc gatcagcctg aagtgccagg atcacctgtt 1800
87 cctcttccgc attaccagcg accggccgct ggaggagctc tttctcgagc agctggaggc 1860
89 gccgccgcca cccggcctgg cgatgaaact ggagtagggt cccgactcta aagtctcccc 1920
91 cgttctccat ccgaaaaatg tttcattgtg attgcgtttg tttgcatttc tcctctctat 1980
93 cccttatacc ctacaaaagc cccctaatat tacgcaaaat gtgtatgtaa ttgtttattt 2040
95 tttttttatt acctaatatt attattatta ttgatataga aaatgttttc cttaagatga 2100
97 agattagcct cctcgacgtt tatgtcccag taaacgaaaa acaaacaaaa tccaaaactt 2160
99 gaaaagaaca caaaacacga acgagaaaat gcacacaagc aaagtaaaag taaaagttaa 2220
101 actaaagcta aacgagtaaa gatattaaaa taacggttaa aattaatgca tagttatgat 2280
103 ctacagacgt atgtaaacat acaaattcag cataaatata tatgtcagca ggcgcatatc 2340
105 tgcggtgctg gccccgttct aaatcaattg taattacttt ttaacataaa tttacccaaa 2400
107 acgttatcaa ttagatgcga gatacaaaaa tcaccgacga aaaccaacaa aatatatcta 2460
109 tgtataaaaa atataaactg cataacaa 2488
112 <210> SEQ ID NO: 2
113 <211> LENGTH: 508
114 <212> TYPE: PRT
115 <213> ORGANISM: Drosophila melanogaster
117 <400> SEQUENCE: 2
119 Met Asp Asn Cys Asp Gln Asp Ala Ser Phe Arg Leu Ser His Ile Lys
120 1 5 10 15
123 Glu Glu Val Lys Pro Asp Ile Ser Gln Leu Asn Asp Ser Asn Asn Ser
124 20 25 30
127 Ser Phe Ser Pro Lys Ala Glu Ser Pro Val Pro Phe Met Gln Ala Met
128 35 40 45
131 Ser Met Val His Val Leu Pro Gly Ser Asn Ser Ala Ser Ser Asn Asn
132 50 55 60
135 Asn Ser Ala Gly Asp Ala Gln Met Ala Gln Ala Pro Asn Ser Ala Gly
136 65 70 75 80
139 Gly Ser Ala Ala Ala Ala Val Gln Gln Gln Tyr Pro Pro Asn His Pro
140 85 90 95
143 Leu Ser Gly Ser Lys His Leu Cys Ser Ile Cys Gly Asp Arg Ala Ser
144 100 105 110
147 Gly Lys His Tyr Gly Val Tyr Ser Cys Glu Gly Cys Lys Gly Phe Phe
148 115 120 125
151 Lys Arg Thr Val Arg Lys Asp Leu Thr Tyr Ala Cys Arg Glu Asn Arg
152 130 135 140
155 Asn Cys Ile Ile Asp Lys Arg Gln Arg Asn Arg Cys Gln Tyr Cys Arg
156 145 150 155 160
159 Tyr Gln Lys Cys Leu Thr Cys Gly Met Lys Arg Glu Ala Val Gln Glu
160 165 170 175
163 Glu Arg Gln Arg Gly Ala Arg Asn Ala Ala Gly Arg Leu Ser Ala Ser
164 180 185 190
167 Gly Gly Gly Ser Ser Gly Pro Gly Ser Val Gly Gly Ser Ser Ser Gln
168 195 200 205
171 Gly Gly Gly Gly Gly Gly Gly Val Ser Gly Gly Met Gly Ser Gly Asn
172 210 215 220
175 Gly Ser Asp Asp Phe Met Thr Asn Ser Val Ser Arg Asp Phe Ser Ile

RAW SEQUENCE LISTING DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:37

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

```
176 225                 230                 235                 240
179 Glu Arg Ile Ile Glu Ala Glu Gln Arg Ala Glu Thr Gln Cys Gly Asp
180                 245                 250                 255
183 Arg Ala Leu Thr Phe Leu Arg Val Gly Pro Tyr Ser Thr Val Gln Pro
184             260                 265                 270
187 Asp Tyr Lys Gly Ala Val Ser Ala Leu Cys Gln Val Val Asn Lys Gln
188         275                 280                 285
191 Leu Phe Gln Met Val Glu Tyr Ala Arg Met Met Pro His Phe Ala Gln
192     290                 295                 300
195 Val Pro Leu Asp Asp Gln Val Ile Leu Leu Lys Ala Ala Trp Ile Glu
196 305                 310                 315                 320
199 Leu Leu Ile Ala Asn Val Ala Trp Cys Ser Ile Val Ser Leu Asp Asp
200                 325                 330                 335
203 Gly Gly Ala Gly Gly Gly Gly Gly Gly Leu Gly His Asp Gly Ser Phe
204             340                 345                 350
207 Glu Arg Arg Ser Pro Gly Leu Gln Pro Gln Gln Leu Phe Leu Asn Gln
208         355                 360                 365
211 Ser Phe Ser Tyr His Arg Asn Ser Ala Ile Lys Ala Gly Val Ser Ala
212     370                 375                 380
215 Ile Phe Asp Arg Ile Leu Ser Glu Leu Ser Val Lys Met Lys Arg Leu
216 385                 390                 395                 400
219 Asn Leu Asp Arg Arg Glu Leu Ser Cys Leu Lys Ala Ile Ile Leu Tyr
220                 405                 410                 415
223 Asn Pro Asp Ile Arg Gly Ile Lys Ser Arg Ala Glu Ile Glu Met Cys
224             420                 425                 430
227 Arg Glu Lys Val Tyr Ala Cys Leu Asp Glu His Cys Arg Leu Glu His
228         435                 440                 445
231 Pro Gly Asp Asp Gly Arg Phe Ala Gln Leu Leu Leu Arg Leu Pro Ala
232     450                 455                 460
235 Leu Arg Ser Ile Ser Leu Lys Cys Gln Asp His Leu Phe Leu Phe Arg
236 465                 470                 475                 480
239 Ile Thr Ser Asp Arg Pro Leu Glu Glu Leu Phe Leu Glu Gln Leu Glu
240                 485                 490                 495
243 Ala Pro Pro Pro Pro Gly Leu Ala Met Lys Leu Glu
244             500                 505
247 <210> SEQ ID NO: 3
248 <211> LENGTH: 61
249 <212> TYPE: DNA
250 <213> ORGANISM: Trichoplusia ni granulovirus
252 <400> SEQUENCE: 3
253 gaccaattaa taggtgacct gcgataaaaa ttacctataa atatgtgatg ttgctggatt     60
255 g                                                                    61
258 <210> SEQ ID NO: 4
259 <211> LENGTH: 134
260 <212> TYPE: DNA
261 <213> ORGANISM: Trichoplusia ni granulovirus
263 <400> SEQUENCE: 4
264 cgagaggtta tcgcccaata caacaacaat gataatgacg tgcaagcaga taatagtgaa     60
266 aaaataacag atactagagt ataaaaaggg gatgctggga gtggacaggc acagtcgtgg    120
```

RAW SEQUENCE LISTING DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:37

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

```
268 tgtggcagca aaca                                                    134
271 <210> SEQ ID NO: 5
272 <211> LENGTH: 69
273 <212> TYPE: DNA
274 <213> ORGANISM: Trichoplusia ni granulovirus
276 <400> SEQUENCE: 5
277 tcagtataaa aaggggtgca ttctcggtaa gagtacagtt gaactcacat cgagttaact     60
279 ccacgatga                                                          69
282 <210> SEQ ID NO: 6
283 <211> LENGTH: 63
284 <212> TYPE: DNA
285 <213> ORGANISM: Trichoplusia ni granulovirus
287 <400> SEQUENCE: 6
288 taagggtagt ataaaaaggc gatcaatcat tgacaaacag tttgcagcag gctgtgggaa     60
290 cga                                                                63
293 <210> SEQ ID NO: 7
294 <211> LENGTH: 15
295 <212> TYPE: DNA
296 <213> ORGANISM: Artificial Sequence
298 <220> FEATURE:
299 <223> OTHER INFORMATION: Chemically synthesized
301 <400> SEQUENCE: 7
302 gaggtcaatg acctc                                                     15
305 <210> SEQ ID NO: 8
306 <211> LENGTH: 24
307 <212> TYPE: DNA
308 <213> ORGANISM: Artificial Sequence
310 <220> FEATURE:
311 <223> OTHER INFORMATION: Chemically synthesized
314 <220> FEATURE:
315 <221> NAME/KEY: misc_feature
316 <222> LOCATION: (7)..(18)
317 <223> OTHER INFORMATION: N is A, T, G or C
319 <220> FEATURE:
320 <221> NAME/KEY: misc_feature
321 <222> LOCATION: (8)..(18)
322 <223> OTHER INFORMATION: Any one of these 11 N's may or may not be present
324 <400> SEQUENCE: 8
W--> 325 aggtcannnn nnnnnnnnag gtca                                      24
328 <210> SEQ ID NO: 9
329 <211> LENGTH: 24
330 <212> TYPE: DNA
331 <213> ORGANISM: Artificial Sequence
333 <220> FEATURE:
334 <223> OTHER INFORMATION: Chemically synthesized
337 <220> FEATURE:
338 <221> NAME/KEY: misc_feature
339 <222> LOCATION: (7)..(18)
340 <223> OTHER INFORMATION: N = A, T, G or C
```

RAW SEQUENCE LISTING DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:37

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

342 <220> FEATURE:
343 <221> NAME/KEY: misc_feature
344 <222> LOCATION: (8)..(18)
345 <223> OTHER INFORMATION: Any one of these 11 N's may or may not be present
347 <400> SEQUENCE: 9
W--> 348 tgacctnnnn nnnnnnnntg acct **24**
351 <210> SEQ ID NO: 10
352 <211> LENGTH: 13
353 <212> TYPE: DNA
354 <213> ORGANISM: Artificial Sequence
356 <220> FEATURE:
357 <223> OTHER INFORMATION: Chemically synthesized
360 <220> FEATURE:
361 <221> NAME/KEY: misc_feature
362 <222> LOCATION: (7)..(7)
363 <223> OTHER INFORMATION: n is a, c, g, or t
365 <400> SEQUENCE: 10
W--> 366 aggtcanagg tca **13**
369 <210> SEQ ID NO: 11
370 <211> LENGTH: 26
371 <212> TYPE: DNA
372 <213> ORGANISM: Artificial Sequence
374 <220> FEATURE:
375 <223> OTHER INFORMATION: Chemically synthesized
378 <220> FEATURE:
379 <221> NAME/KEY: misc_feature
380 <222> LOCATION: (7)..(7)
381 <223> OTHER INFORMATION: n is a, c, g, or t
383 <220> FEATURE:
384 <221> NAME/KEY: misc_feature
385 <222> LOCATION: (20)..(20)
386 <223> OTHER INFORMATION: n is a, c, g, or t
388 <400> SEQUENCE: 11
W--> 389 aggtcanagg tcaaggtcan aggtca **26**
392 <210> SEQ ID NO: 12
393 <211> LENGTH: 26
394 <212> TYPE: DNA
395 <213> ORGANISM: Artificial Sequence
397 <220> FEATURE:
398 <223> OTHER INFORMATION: Chemically synthesized
401 <220> FEATURE:
402 <221> NAME/KEY: misc_feature
403 <222> LOCATION: (7)..(7)
404 <223> OTHER INFORMATION: n is a, c, g, or t
406 <220> FEATURE:
407 <221> NAME/KEY: misc_feature
408 <222> LOCATION: (20)..(20)
409 <223> OTHER INFORMATION: n is a, c, g, or t
411 <400> SEQUENCE: 12

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:38

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:8; N Pos. 7,8,9,10,11,12,13,14,15,16,17,18
Seq#:9; N Pos. 7,8,9,10,11,12,13,14,15,16,17,18
Seq#:10; N Pos. 7
Seq#:11; N Pos. 7,20
Seq#:12; N Pos. 7,20
Seq#:16; N Pos. 9,10,11,12,13,14,15,16,17,18,19,20

VERIFICATION SUMMARY DATE: 11/15/2004
PATENT APPLICATION: **US/10/719,024** TIME: 15:02:38

Input Set : **D:\US Utility 50229-420 Sequence Listing.txt**
Output Set: **N:\CRF4\11152004\J719024.raw**

L:325 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:348 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:0
L:366 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10 after pos.:0
L:389 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:412 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:0
L:466 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0